DOCUMENT RESUME

ED 074 050 SP 006 276

AUTHOR Frye, Helen B.
TITLE Development and Initial Validation of the Teaching Situation Reaction Test (Elementary Form), A Measure of Teacher Effectiveness.
SPONS AGENCY Dayton Univ., Ohio. Research Inst.
PUB DATE [72]
NOTE 11p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Effective Teaching; *Elementary School Teachers; Measurement Instruments; *Teacher Evaluation; Teaching Quality; *Test Construction; Test Results; *Test Validity
IDENTIFIERS *Teaching Situation Reaction Test; TSRT

ABSTRACT

The Teaching Situation Reaction Test (TSRT), a measure of secondary school teachers' effectiveness, was adapted to the intermediate elementary grades. The TSRT (Elementary Form) was administered to 210 volunteer teachers in the fourth, fifth, and sixth grades in Dayton, Ohio during January 1969. Two groups of teachers were formed for in-depth study--those scoring in the top 27 percent (High Group) and the bottom 27 percent (Low Group) of the 180 teachers providing useable response sets. Each group contained 48 teachers. Teachers in the High Group were compared with teachers in the Low Group on 25 factors in four categories: personal, classroom, school environment, and performance. A one-way analysis of variance was made by computer, comparing the means of the two groups of teachers on all 25 factors. The F-test was applied to determine significant differences. Results indicated that teachers scoring high on the TRST and teachers scoring low represent different samples of the population. Further conclusions are indicated. (A 9-item bibliography and related research material are included.) (Author/MJM)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 074050

SP 006 276

DEVELOPMENT AND INITIAL VALIDATION OF THE TEACHING SITUATION REACTION TEST (ELEMENTARY FORM), A MEASURE OF TEACHER EFFECTIVENESS[1]

One of the major problems concerning effectiveness of in-service teachers and the prediction of effectiveness during the pre-service phase is that of selecting appropriate data on which to base an evaluation. Research, which reflects the difference of opinion, has revealed an interesting fact - individual measures, such as intelligence tests and the Minnesota Multiphasic Personality Inventory, have shown varying degrees of correlation with the same, as well as with different, criterion measures. As early as the 1940's there was a strong indication that different criteria measure different things (2). This is not surprising, however, when one realizes that teacher effectiveness is most likely influenced by more than a single factor, and that a single trait or instrument reflects a very narrow range of behavior. Measures used in combination have given somewhat more consistent results. For example, Simun and Asher found that a combination of nine variables in undergraduate school had a significant correlation with administrators' ratings of teaching ability (8). But use of several measures is not practical.

The problem becomes one of finding a broadly based measure which will correlate with accepted criteria of effectiveness. This implies a paper-and-pencil instrument which will be a quick, indirect measure of teacher behavior as well as of the product of teaching. Both administrators' ratings and

---

[1]This research was made possible by a grant from the University of Dayton Research Council.

measured pupil change are considered appropriate criteria; ratings have been generally accepted for many years, although recently emphasis has been shifting to measurable outcomes of education (1).

A few attempts have been made to develop such broad instruments. One, the Minnesota Teacher Attitude Inventory, has predictive validity for the rather narrow range of behavior it measures, the ability to maintain harmonious relationships with pupils. Inconsistent results have been obtained, however, in research using other effectiveness criteria (4). Another, Turner's set of simulated teaching tasks, is limited to specific subject-matter areas (9).

The Teaching Situation Reaction Test (TSRT) belongs in this category of instruments designed especially for measuring teacher effectiveness. It was originally developed to help with the study of the effectiveness of pre-service secondary education courses. Oriented toward junior-high-school level, the TSRT is comprised of forty-eight sequential problem situations, each with four suggested solutions which must be ranked in order of desirability. The situations are subject-matter neutral, and concern common aspects of teaching such as planning, classroom management, and teacher-pupil relationships. The TSRT was built on the idea that objectivity, empathy, control, confidence, and creativity are essential factors in effective responses to classroom situations (7:1-3). A high score is believed to indicate effective teaching.

The TSRT was developed in 1957 by Duncan and Frymier. Since that time, Amidon and Hough have contributed to its refinement, rekeying, and expansion, and numerous studies of its reliability and validity have been made. Generally speaking, it has been found to have validity for predicting student-teaching grade and to bear a positive relationship to administrators' ratings and pupil change (5:1-12). The fourth revision (September 1966) was used in research conducted by the writer at the sixth-grade level to assess its validity for

elementary school teachers. The results, very briefly, showed a positive relationship between teacher score on the TSRT and pupil change in acceptance of others and grade-point average, but not in the other criterion measures of acceptance of self, motivation toward school, and achievement as measured by standardized test (6:159-70). The research reported herein modified the TSRT for use with elementary teachers of the intermediate grades and studied the relationship between scores on the modified TSRT and 25 other variables presumed to be related.

Modification of TSRT

The study at sixth-grade level referred to previously involved 60 teachers in the final sample. These teachers, originally arranged by TSRT score, were arranged in order according to the mean gain of their class in acceptance of others, and also according to mean class gain in acceptance of self. Each teacher was given a rank-order in each listing and the two ranks averaged. The resulting averages were ordered, and teachers in the top and bottom 27% of the composite rank-order were selected for further study. Their responses to each item on the TSRT were analyzed with the purpose of finding common option-ranking patterns within the high group and within the low group. Comparison of patterns revealed 29 items which discriminated between teachers in the top and bottom 27% (that is, teachers high in promoting pupil acceptance of self and others vs. teachers low in promoting pupil acceptance of self and others). The option ranking for these 29 items as indicated by the top group were used as the basis of a new scoring key.

At this point the assistance of three teachers was enlisted - one in each of the fourth, fifth and sixth grades. They were selected by school officials as being "some of the best" teachers. The 19 items which did not discriminate

between top and bottom were carefully studied by the three teachers and the writer. Some items were rewritten and for some the answer key was altered, always keeping in mind the criterion measures of development of acceptance of self and others in pupils. In addition, other wording changes were made to make the context more appropriate to the intermediate grades. All changes in the TSRT itself were kept to the minimum in order to preserve, as much as possible, the factors inherent in its construction. The remaining instrument was designated as the Teaching Situation Reaction Test (Elementary Form).

Factors Related to Performance on TSRT (Elementary Form)

Hypotheses. This study, part of a comprehensive one concerning both validity and reliability, was designed to determine if teachers scoring high on the TSRT (Elementary Form) do in fact possess characteristics which distinguish them from teachers scoring low, and which are consistent with the construct of the original TSRT. The general hypothesis to be tested was: Teachers who score high on the TSRT (Elementary Form) do not differ significantly from teachers who score low on any of the measures available for this study. Two specific hypotheses related to the construct were also tested: (1) Teachers who score high do not rate significantly higher than teachers who score low on the five performance characteristics related to primary factors built into the original TSRT; and (2) Teachers who score high do not rate significantly higher than teachers who score low on the principal's rating of overall teaching effectiveness.

Procedure. The research design involved examination of the construct validity by employing a variation on the known-groups method.

The TSRT (Elementary Form) was administered to teachers in the fourth, fifth, and sixth grades of the Dayton, Ohio, public schools during January, 1969. Twenty-six of the 50 elementary schools having all three grades were selected to participate, representing a cross-section in terms of size, geographic location, and socio-economic level. A total of 210 volunteer teachers completed the TSRT (Elementary Form), providing 180 useable response sets for this part of the comprehensive study.

Two groups of teachers were formed for in-depth study - those scoring in the top 27% of the 180 teachers (High Group), and those scoring in the bottom 27% (Low Group). Each group contained 48 teachers. A perfect score would be 288 (6 points for each of the 48 questions). The score range in the High Group was 210 to 226, and in the Low Group, 152 to 192. More complete descriptive statistics on the two groups are provided in Table 1.

As much information as possible was collected on the 96 teachers without giving them additional tests. The central office and school principals furnished personal, situational, and performance data on teachers.

Teachers in the High Group were compared with teachers in the Low Group on 25 factors, which can be roughly grouped into 4 categories - personal, classroom, school-environment, and performance. The seven personal factors were sex, age, race, years of teaching experience, master's degree, bachelor's degree, and days absent during the school year. The three classroom factors were size of class, mean class score on pupil Acceptance of Self, and mean class score on pupil Acceptance of Others. (The last two were obtained by using Bill's Index of Adjustment and Values.(3)) The nine school-environment factors were sex of principal, professional attitude of principal (subjective rating by the writer based on cooperation, interest, understanding, etc.), school size - faculty,

school size - pupils, school climate (subjective rating of faculty based on cooperation, friendliness, group cohesiveness, etc.), mean TSRT score in the school, race of pupils, race of faculty, and whether or not school was classified as "disadvantaged" (determined by school system on basis of state and federal aid, families on ADC, etc.). Factors in the performance category were taken from a rating sheet completed by the teacher's principal - social sensitivity, direct control in the classroom, organization, emotional stability, staff relationships, and over-all teaching effectiveness. The first five factors were selected as closely representing those supposedly built into the original TSRT.

Findings and Discussion. A simple one-way analysis of variance was made by computer, comparing the means of the two groups of teachers on all 25 factors, to discover potential relationships among variables. The F-test was applied to determine significance of differences. Table 2 shows the results of analysis of variance; 14 factors discriminated between the High Group and the Low Group. All discriminations shown favored the High Group (higher scores, ratings, percentages, etc.).

Only one personal factor (race) distinguished between teachers in the High and Low Groups, with a significantly higher proportion of white teachers in the High Group than in the Low Group. When this is combined with the fact that the High Group tended to teach in the schools having a significantly higher percentage of both white pupils and white faculty members, two possible explanations seem worth considering. One, there are cultural differences which influence teachers' reactions to classroom situations. Two, schools with a relatively high percentage of black pupils and faculty have a different influence on a teacher's attitudes toward classroom situations than do schools with

white students and faculty. (The influence would be a detrimental or negative one if a high TSRT score truly indicates an effective teacher.)

The only classroom factor which discriminated was class size - High Group teachers had larger classes. This may not be surprising when one considers that principals rated High Group teachers as significantly more "effective" than Low Group teachers. Is it not logical to assign more students to teachers who are "effective"?

Six of the nine school-environment factors showed a significant difference between the two teacher groups, those related to race, poverty, and school climate. There is a strong relationship between the school environment and teacher score, however, it is impossible to tell from the data available if a cause-effect relationship exists or, if it does, in which direction it operates.

All six performance factors discriminated in favor of the High Group. Consistent significant differences on the first five indicate strongly that the construct of the original secondary-school TSRT has been preserved in this modification for the elementary school. The highly significant difference on the final factor, "Over-all Effectiveness," suggests that the TSRT (Elementary Form) is a valid measure of teacher effectiveness when principal's ratings are the criterion of performance. An interesting point to be noted is that the difference is significant at the .01 level, as compared with the .05 level for all of the five factors believed to contribute to effectiveness. There is likely another factor(s) involved which contributes to total effectiveness in a greater degree than the five listed. Or, possibly, it is easier for a principal to assess a teacher's performance globally rather than by component parts, if for no other reason than the lack of universal definitions of the terms used to identify the parts.

On the basis of the findings, the general hypothesis must be partially rejected and the other two hypotheses must be completely rejected.

Conclusions

The following conclusions seem warranted, keeping in mind that the research was done in a large metropolitan school system and participation in the project was voluntary.

a. Teachers scoring high on the TSRT and teachers scoring low represent different samples of the population, as evidenced by significant differences on 14 of the 25 factors examined.

b. The underlying construct of the original TSRT has apparently not been appreciably altered in this revision for the intermediate elementary grades, in that teachers scoring high on the TSRT (Elementary Form) are considered by their principals to possess the qualities built into the original TSRT to a significantly higher degree than teachers scoring low.

c. Conditions of the school environment having to do with race, climate, and poverty, and teacher effectiveness as indicated by principal's rating, are related, as both are significantly related to teacher score on the TSRT (Elementary Form). Additional research is needed to determine the nature of the relationship. The data suggest a definite educational justification for the current move to desegregate schools and work toward a racial balance of both teachers and pupils.

d. The TSRT (Elementary Form) is a valid indicator of teacher effectiveness in the intermediate elementary grades when principal's rating is used as the criterion measure. Additional research is needed to asses its validity when measured pupil change is the criterion of effectiveness.

References

1. Ackerman, Walter I.,"Teacher Competence and Pupil Change," Harvard Educational Review, 24: 273-89, 1954.

2. Barr, A. S., "The Measurement and Prediction of Teaching Efficiency: A Summary of Investigations," Journal of Experimental Education, 16: 203-83, 1948.

3. Bills, Robert E., "Manual for Index of Adjustment and Values."

4. Cook, Walter W., et al., "Significant Factors in Teachers' Classroom Attitudes," Journal of Teacher Education, 7: 274-80, 1956.

5. Duncan, James K., and Hough, John B., "Technical Review of the Teaching Situation Reaction Test," The Ohio State University, September, 1966.

6. Frye, Helen B., "A Study of the Teaching Situation Reaction Test in Relation to Effectiveness of In-Service Teachers," Ph.D. dissertation, The Ohio State University, 1967.

7. Hough, John B., and Duncan, James K., "Exploratory Studies of a Teaching Situation Reaction Test," paper read at the American Educational Research Association, Chicago, Illinois, February, 1965.

8. Simun, Patricia Bates; and Asher, John William, "The Relationship of Variables in Undergraduate School and School Administrators' Ratings of First-Year Teachers," Journal of Teacher Education, 15: 293-302, 1964.

9. Turner, Richard L., "Task Performance and Teaching Skill in the Intermediate Grades," Journal of Teacher Education, 14: 299-307, 1963.

TABLE I

Data on TSRT Scores of High and Low Teacher Groups

ITEM	HIGH GROUP	LOW GROUP
Number	48	48
Range	210-226	152-192
Mean	216.44	182.38
Standard Deviation	4.72	10.37
Standard Error of Mean	.68	1.50

Table 2

Comparison of Teachers in High Group and Low Group on 25 Factors Believed to be Related to Teaching Effectiveness

Factor	F Ratio	p*
Personal		
Sex	0.4393	n.s.
Age	0.7729	n.s.
Race (white, Black)	8.0764	.01
Years of Experience	0.0748	n.s.
Master's Degree	0.3287	n.s.
Bachelor's Degree	0.5371	n.s.
Days Absent	0.8836	n.s.
Classroom		
Class Size	4.6006	.05
Mean Pupil Score, Acceptance of Self	0.0092	n.s.
Mean Pupil Score, Acceptance of Others	1.7218	n.s.
School Environment		
Principal's Professional Attitude	9.6720	.01
Sex of Principal	0.5818	n.s.
School Climate	6.9030	<.05
School Size - Faculty	1.1354	n.s.
School Size - Pupils	0.2225	n.s.
Mean TSRT Score in School	7.2763	.01
Pupil Race (% white in school)	5.7364	.05
Faculty Race (% white in school)	8.1808	.01
Advantaged/Disadvantaged Population	16.2000	<.01
Performance		
Social Sensitivity	5.0349	.05
Direct Control	4.1191	.05
Organization	5.2132	.05
Emotional Stability	6.9311	<.05
Staff Relationships	5.9407	.05
Over-all Effectiveness	7.6429	.01

*For last 6 factors, entering the F table with 1 df between groups and 86 df within groups, $F_{.05}=3.95$ and $F_{.01}=6.94$. For all other factors, entering the F table with 1 df between groups and 94 df within groups, $F_{.05}=3.95$ and $F_{.01}=6.91$.